1 Supplementary files for

Inflammatory microbes and genes as potential biomarkers of Parkinson's disease

4

5 Shiqing Nie[1,2], Jichen Wang[1,2], Ye Deng[2,3], Zheng Ye[4], Yuan Ge[1,2,*]

6

7 [1]State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental

8 Sciences, Chinese Academy of Sciences, Beijing 100085, China

9 [2]University of Chinese Academy of Sciences, 100049 Beijing, China

10 [3]Key Laboratory of Environmental Biotechnology, Research Center for Eco-Environmental

11 Science, Chinese Academy of Sciences, Beijing 100085, China

12 [4]Institute of Neuroscience, Center for Excellence in Brain Science and Intelligence Technology,

13 Chinese Academy of Sciences, Shanghai 200031, China

14

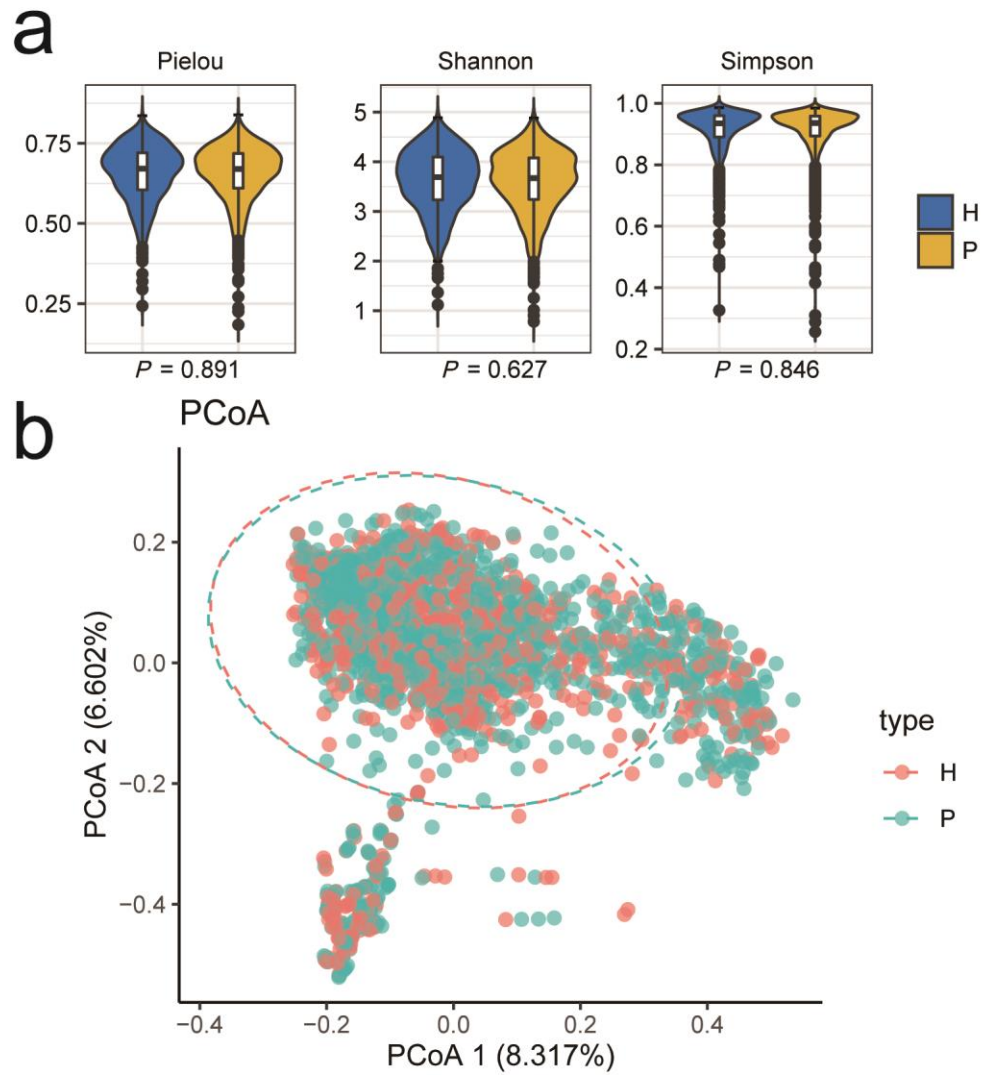15 [*]To whom correspondence may be addressed. Email: yuange@rcees.ac.cn

16

Competing interests
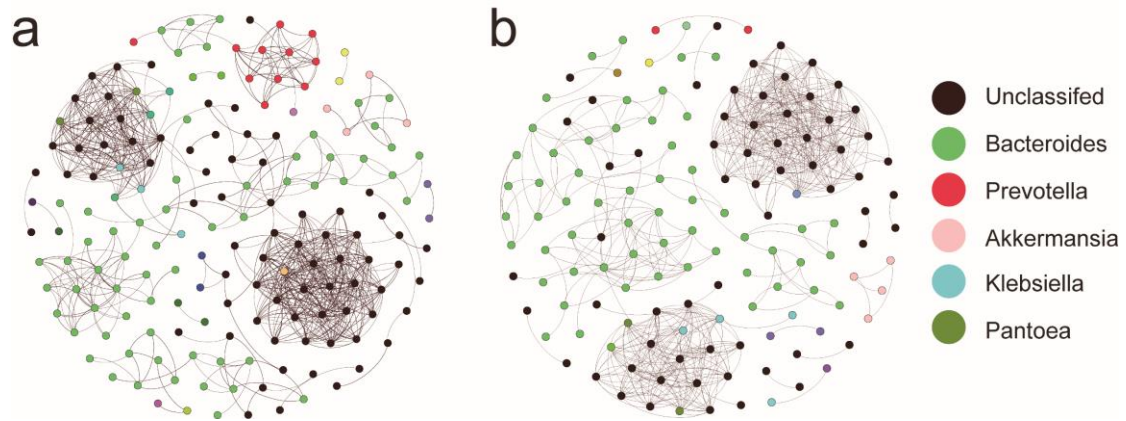
The authors declare no competing interests.

19

20

21

Supplementary Figure 1: The comparison of microbial diversity between Parkinson's disease patients and healthy controls. (a) Alpha diversity indexes. (b) Principal coordinates analysis (PCoA) based on Bray-Curtis distance. H (Healthy control). P (Parkinson's disease patient).

26

27

Supplementary Figure 2: Co-occurrence networks. Co-occurrence networks were calculated based on Spearman correlations at the operational taxonomic unit (OTU) level. Only strong ($r > 0.7$) and significant ($p < 0.05$) correlations are shown in the figure. (a) Healthy controls. (b) Patients with Parkinson's disease (PD). The difference between this figure and Fig. 5 is that the nodes with few connections (< 5) were not removed and the nodes in this figure were colored by genus. On the right of the graph, only the unclassified and five genera with the highest number of nodes were labeled.
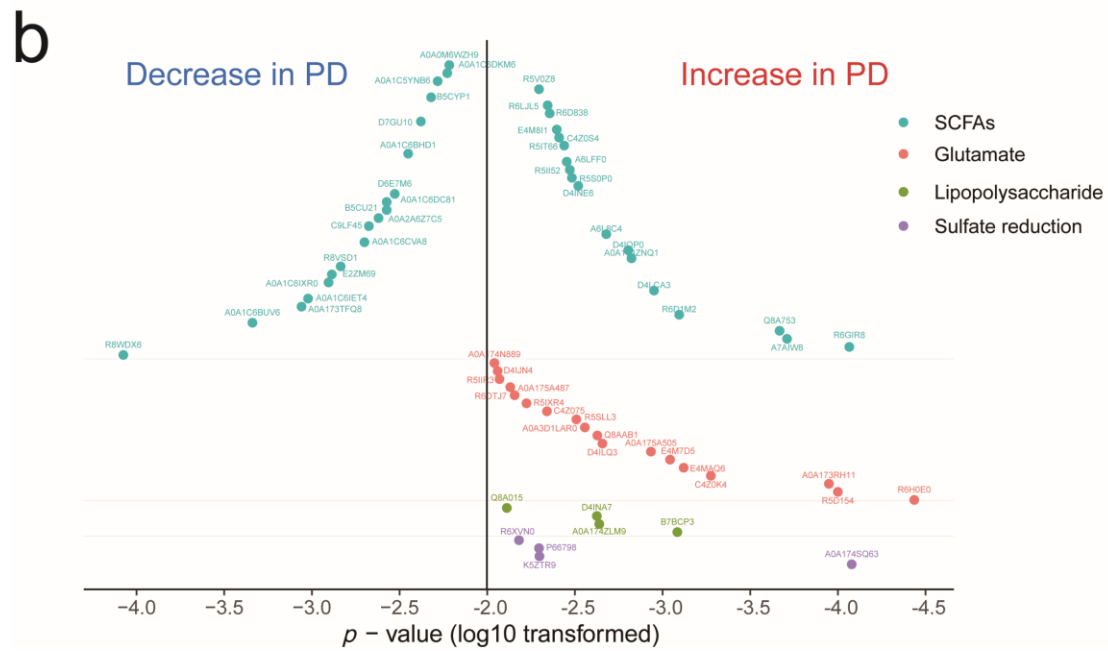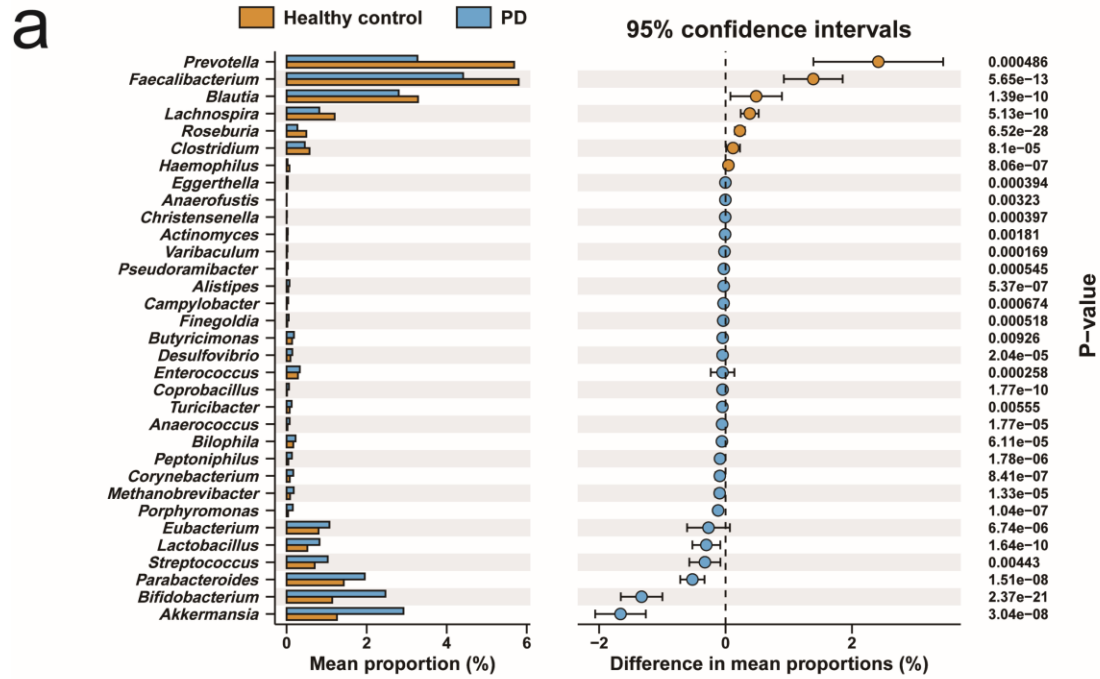
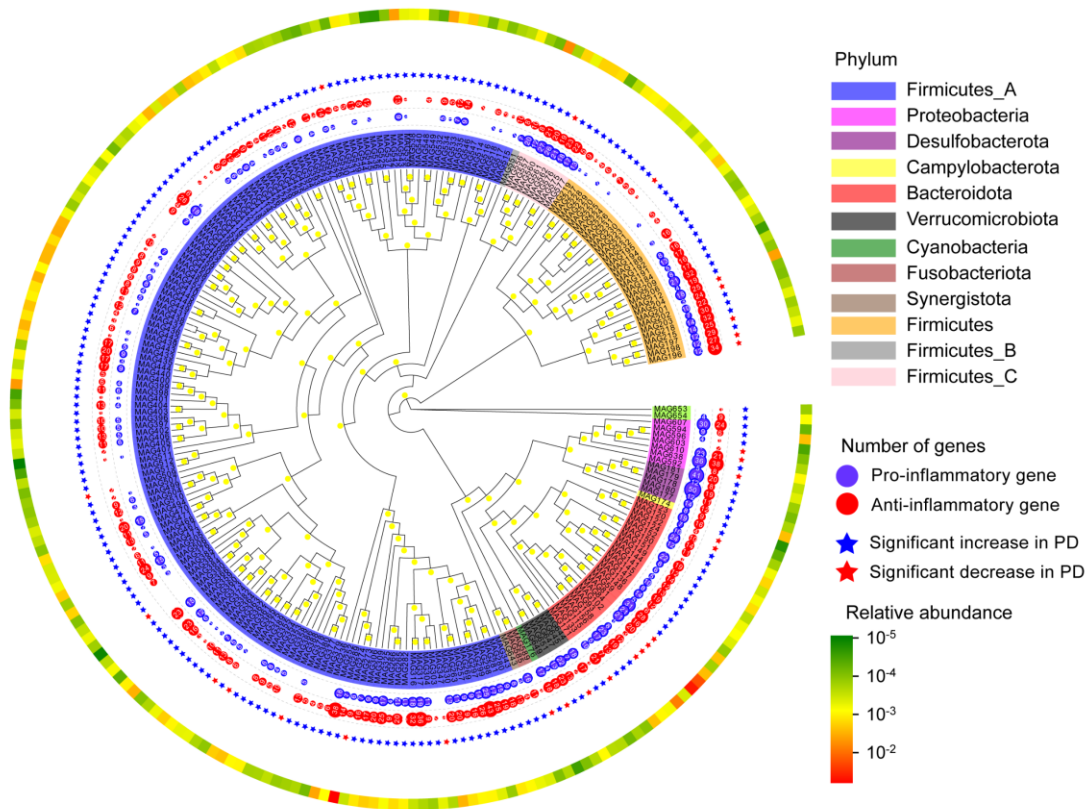**Supplementary Figure 3: Differences in microorganisms and genes between healthy controls and patients with Parkinson's disease (PD).** (a) Genera with significant differences in relative abundance. (b) Genes (related to inflammation) with significant differences in transcripts per million (TPM) abundance. SCFAs: short-chain fatty acids.

44

Supplementary Figure 4: The maximum-likelihood phylogenetic tree. The tree was based on the 120 bacterial concatenated ribosomal proteins, colored by phyla. Bootstrap values were calculated based on 1000 replicates, and the value higher than 80% were marked in yellow. The size of the blue (pro-inflammatory gene) and red (anti-inflammatory gene) circle indicates the number of inflammation-related genes. Blue (significant increase in PD) and red (significant decrease in PD) stars indicate significant differences in the relative abundance of metagenome-assembled genomes (MAGs). Parkinson's disease (PD).
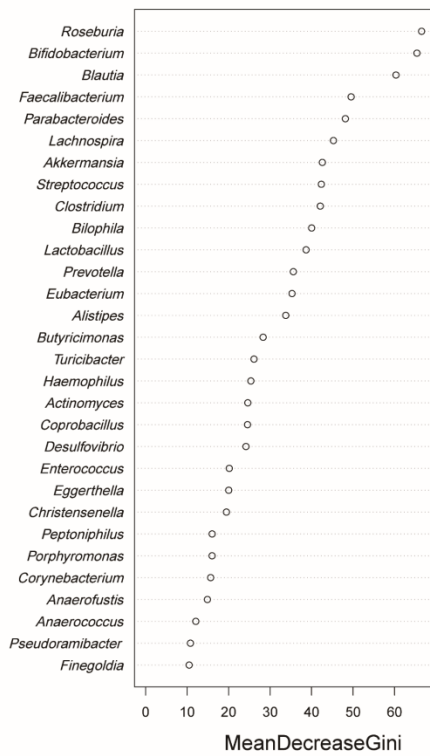
52

53

54
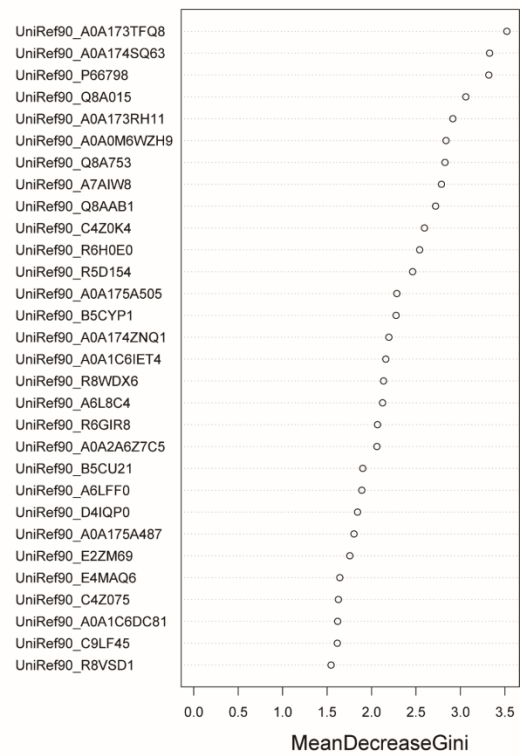
Supplementary Figure 5: The importance of variables in the random forests model. (a) 32 genera. (b) 63 genes. The MeanDecreaseGini values were calculated by the R package "randomForest".

60

Supplementary Figure 6: Determination of the optimal number of variables in the random forests model. The optimal random forest models can be obtained when selecting the 11 genera or the 6 genes.

61

62

63

64

a Variable Importance Plot − PSA Score

	MeanDecreaseGini
Bifidobacterium	
Roseburia	
Blautia	
Faecalibacterium	
Parabacteroides	
Lachnospira	
Clostridium	
Streptococcus	
Akkermansia	
Lactobacillus	
Bilophila	

0   20   40   60   80   100
MeanDecreaseGini

b Variable Importance Plot − PSA Score

UniRef90_A0A174SQ63 (Thioredoxin) — Sulfate reduction (trxA)
UniRef90_Q8A015 (Bifunctional enzyme) — Lipopolysaccharide biosynthesis (LpxC/FabZ)
UniRef90_A0A173RH11 (Citrate synthase) — SCFAs (gltA)
UniRef90_A0A0M6WZH9 (Glutamate synthase) — Glutamate biosynthesis (gltD)
UniRef90_A0A173TFQ (Fructose-1,6-bisphosphate aldolase) — SCFAs (kbaY)
UniRef90_P66798 (Response regulator) — Sulfate reduction (uvrY)
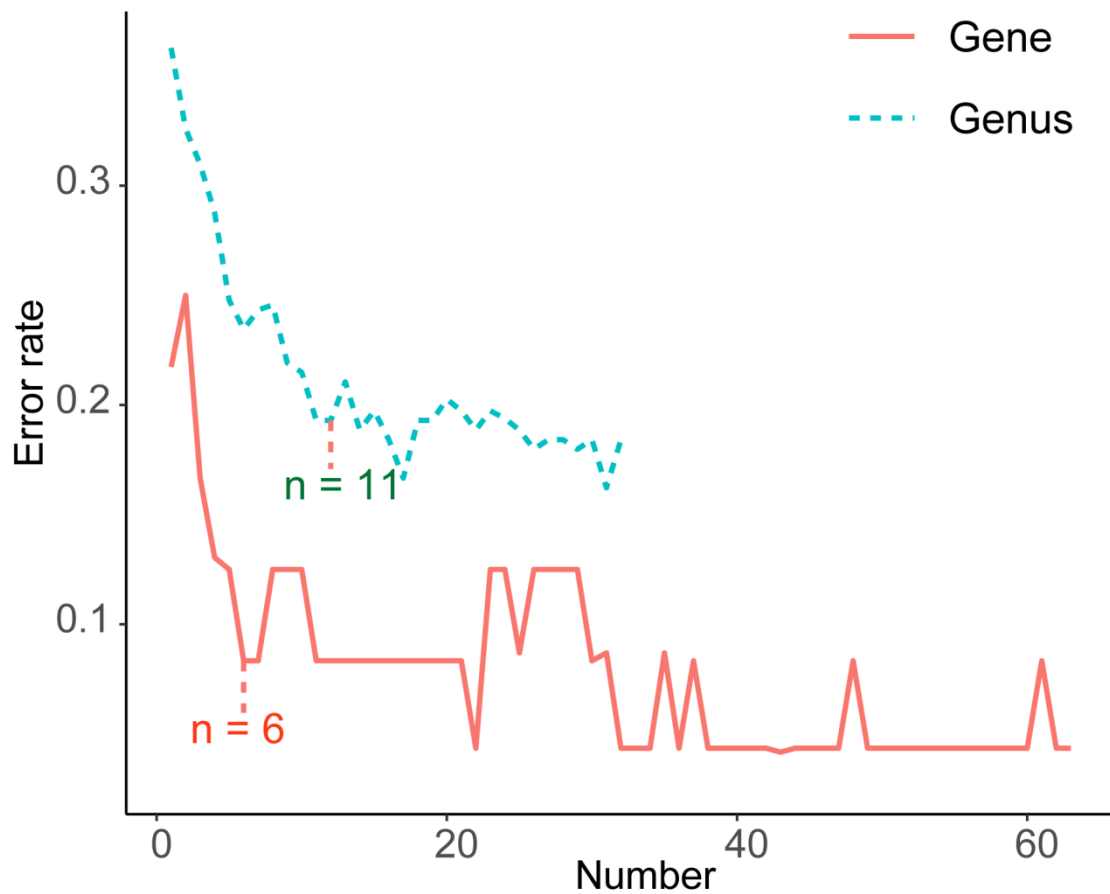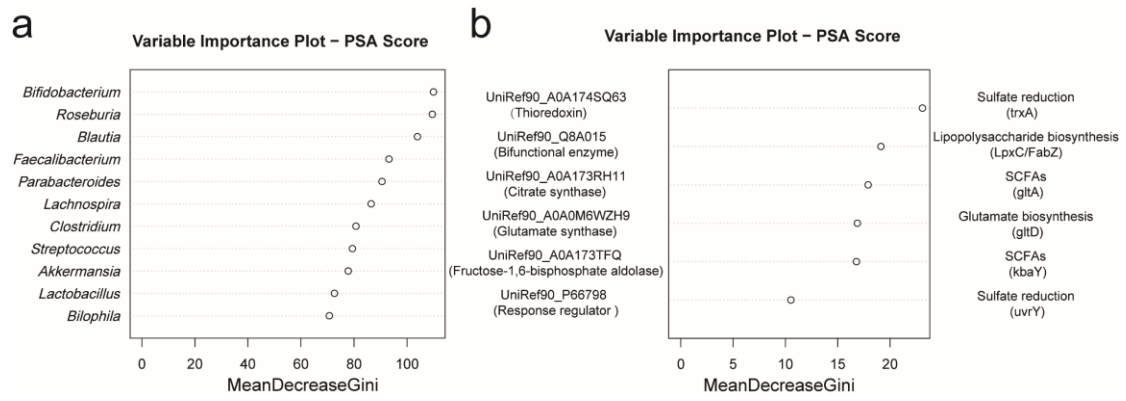
0   5   10   15   20
MeanDecreaseGini

Supplementary Figure 7: The importance of variables in the random forests model. (a) 11 genera. (b) 6 genes. The MeanDecreaseGini values were calculated by the R package "randomForest". SCFAs: short-chain fatty acids.